

# SATB1 reprogrammes gene expression to promote breast tumour growth and metastasis

Hye-Jung Han<sup>1</sup>, Jose Russo<sup>2</sup>, Yoshinori Kohwi<sup>1\*</sup> & Terumi Kohwi-Shigematsu<sup>1\*</sup>

**Mechanisms underlying global changes in gene expression during tumour progression are poorly understood. SATB1 is a genome organizer that tethers multiple genomic loci and recruits chromatin-remodelling enzymes to regulate chromatin structure and gene expression. Here we show that SATB1 is expressed by aggressive breast cancer cells and its expression level has high prognostic significance ( $P < 0.0001$ ), independent of lymph-node status. RNA-interference-mediated knockdown of SATB1 in highly aggressive (MDA-MB-231) cancer cells altered the expression of >1,000 genes, reversing tumorigenesis by restoring breast-like acinar polarity and inhibiting tumour growth and metastasis *in vivo*. Conversely, ectopic SATB1 expression in non-aggressive (SKBR3) cells led to gene expression patterns consistent with aggressive-tumour phenotypes, acquiring metastatic activity *in vivo*. SATB1 delineates specific epigenetic modifications at target gene loci, directly upregulating metastasis-associated genes while downregulating tumour-suppressor genes. SATB1 reprogrammes chromatin organization and the transcription profiles of breast tumours to promote growth and metastasis; this is a new mechanism of tumour progression.**

Metastasis is the final step in solid tumour progression and is the most common cause of death in cancer patients<sup>1</sup>. Metastasis is a multi-step process: invasion of tumour cells into the adjacent tissues, entry of tumour cells in the systemic circulation (intravasation), survival in circulation, extravasation to distant organs, and finally growth of cancer cells to produce secondary tumours<sup>2,3</sup>. How tumour cells become metastatic is largely unknown. It was widely believed that metastatic cells are rare and evolve during late stages of tumour progression from a series of genetic changes that enable the cells to progress through the sequential steps that finally result in growth in distant organ microenvironments. Recently, however, gene expression analysis of human breast carcinomas with known clinical outcomes has revealed profiles that are associated with disease progression and has identified groups of genes whose characteristic expression pattern can predict the risk of metastatic recurrence<sup>4–9</sup>. The detection in some primary tumours of such poor-prognosis gene ‘signatures’ indicates that a large number of cells in the primary tumours already have such a gene expression pattern. Therefore, in addition to the traditional view of metastasis as an evolving process of rare variant clones, the poor-prognosis signatures suggest that cells in some primary tumours are predisposed to metastasis<sup>10</sup>. In fact, there might be an active molecular mechanism underlying such events. How gene expression profiles are established in these tumour cells such that they acquire metastatic properties is unknown.

Here we show that the protein SATB1 is necessary for breast cancer cells to become metastatic, and when ectopically expressed in non-metastatic cells, can induce invasive activity *in vivo*. We also show that SATB1 expression in breast cancer cells establishes gene expression profiles consistent with invasive tumours. SATB1 is a nuclear protein that functions as a ‘genome organizer’ essential for proper T-cell development<sup>11</sup>. SATB1 constitutes a functional nuclear

architecture that has a ‘cage-like’ protein distribution surrounding heterochromatin. This architecture is referred to as ‘the SATB1 regulatory network’, as SATB1 regulates gene expression<sup>11–14</sup> by recruiting chromatin remodelling/modifying enzymes and transcription factors<sup>13,14</sup> to genomic DNA, which it tethers via specialized DNA sequences highly potentiated for unpairing (base unpairing regions, or BURs)<sup>15–17</sup>. On T-helper 2 cell activation, SATB1 becomes expressed and folds the cytokine-gene locus into dense loops for rapid induction of multiple cytokine genes<sup>18</sup>. In breast cancer cells, we find that once SATB1 is expressed, it coordinates expression of a large number of genes to induce metastasis. Removal of SATB1 from aggressive breast cancer cells not only reverses metastatic phenotypes but also inhibits tumour growth, indicating its key role in breast cancer progression.

## SATB1 expression correlates with poor prognosis

We examined SATB1 expression in 24 breast epithelial cell lines, including normal human mammary epithelial cells (HMECs), 5 immortalized derivatives, 13 non-metastatic cancer cell lines and 5 metastatic cancer cell lines. Both SATB1 messenger RNA and protein were detected only in metastatic cancer cell lines, correlating SATB1 expression with aggressive tumour phenotypes (results from representative cell lines shown in Fig. 1a). SATB2, a close homologue of SATB1, was expressed in both malignant and non-malignant cell lines (Supplementary Fig. 1a).

Among 28 human primary breast tumours, SATB1 protein was detected in all 16 poorly differentiated infiltrating ductal carcinomas ( $P < 0.0001$ ). Low-level SATB1 expression was found in some moderately differentiated tumour samples (7 out of 12), and no SATB1 was detected in control samples taken from adjacent non-malignant tissues (representative data shown in Fig. 1b and Supplementary Table

<sup>1</sup>Life Sciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, California 94720, USA. <sup>2</sup>Breast Cancer Research Laboratory, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, USA.

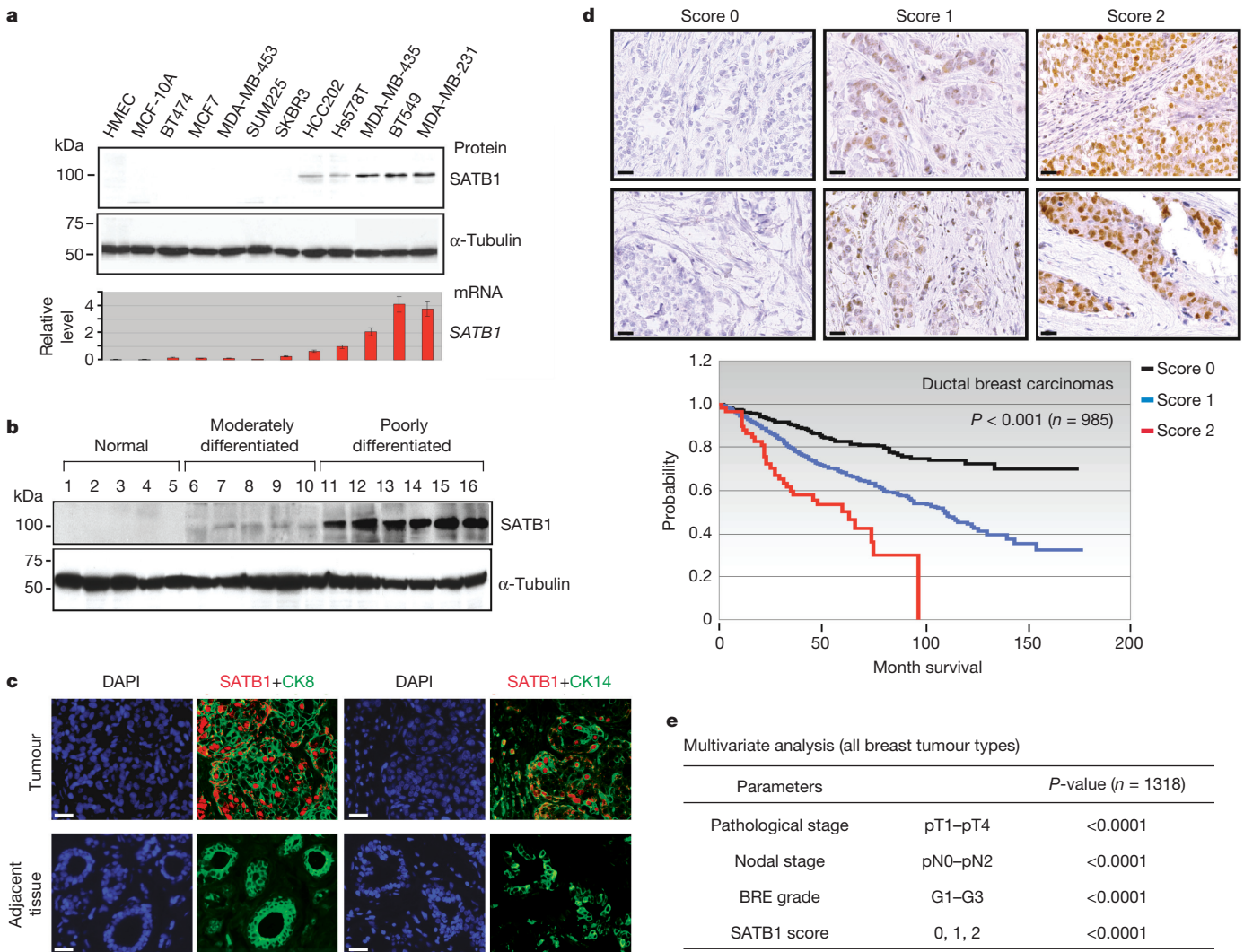
\*These authors contributed equally to this work.

1). Representative immunostaining images of SATB1 and epithelial cell markers in invasive ductal carcinomas are shown in Fig. 1c. The prognostic significance of SATB1 was determined by assessing its nuclear staining using tissue microarrays containing 2,197 cases with known clinical follow-up records, from which 1,318 breast cancer specimens were analysable (Supplementary Table 2 shows tumour composition and SATB1 association with clinico-pathological parameters). Tissues were scored on the basis of the intensity of SATB1 nuclear labelling and percentage of SATB1-positive tumour cells (see Methods). Among these specimens, Kaplan–Meier survival analysis of 985 ductal carcinoma specimens revealed a correlation between higher SATB1 expression levels and shorter overall survival times ( $P < 0.001$ ) (Fig. 1d). This correlation was also observed with all breast cancer types (1,318 specimens) (Supplementary Fig. 1b), except medullary

cancer, which is rare and often has a relatively favourable prognosis despite its poorly differentiated nuclear grade. To exclude the possibility that the prognostic effect of nuclear SATB1 expression was dependent on other established prognostic factors for breast cancer, including tumour stage, the histological grade (Bloom, Richardson, Elston–Ellis grading, BRE) and nodal stage, we performed a multivariate analysis. This analysis confirmed that SATB1 is an independent prognostic factor for breast cancer (Fig. 1e).

SATB1 promotes aggressive cancer phenotypes *in vitro*

We investigated whether SATB1 is required for the invasive phenotypes of breast cancer cells *in vitro* by expressing short hairpin RNAs (shRNA) to knock down SATB1 expression. We expressed shRNA from two different SATB1 sequences (shRNA1 or shRNA2) in the



**Figure 1 | SATB1 expression in breast cancer is associated with poor prognosis.** **a**, Immunoblot analysis of SATB1 levels (top panel) in normal mammary epithelial cells (HMEC), immortalized mammary epithelial cells (MCF-10A), non-aggressive breast cancer cell lines (BT474, MCF7, MDA-MB-435, SUM225 and SKBR3) and aggressive breast cancer cell lines (HCC202, Hs578T, MDA-MB-435, BT549, MDA-MB-231);  $\alpha$ -tubulin loading control is shown in the middle panel. The bottom panels shows transcript levels of SATB1, relative to GAPDH, determined by qRT-PCR and compared to Hs578T cells; error bars indicate s.e.m.,  $n = 3$  experiments. **b**, Immunoblot analysis of SATB1 in representative human primary breast tumour specimen (top panel);  $\alpha$ -tubulin loading control is also shown (bottom panel). **c**, Immunofluorescence images of poorly differentiated ductal carcinomas (top row) and adjacent normal tissues (bottom row)

stained with anti-SATB1 (red) and anti-cytokeratin 8 (CK8) or anti-cytokeratin 14 (CK14) (green) antibodies, counterstained with 4,6-diamidino-2-phenylindole (DAPI, to stain DNA; blue). Scale bars, 30  $\mu$ m. **d**, SATB1 levels in representative tumour tissues (top; scale bar, 20  $\mu$ m) and Kaplan–Meier plot (below) of overall survival of 985 patients with ductal breast carcinomas stratified by SATB1 expression level. Tissues scored as 0 (negative SATB1 nuclear staining for all tumour cells), 1 (positive SATB1 nuclear staining other than score 2) or 2 (moderate SATB1 nuclear staining for >50% tumour cells or strong staining for >5% tumour cells). A log-rank test showed significant differences between groups ( $P < 0.001$ ). **e**, Relative multivariate significance of potential prognostic variables. Cox proportional hazards regression was used to test the independent prognostic contribution of SATB1 after accounting for other potentially important covariates.

highly metastatic MDA-MB-231 cell line. Expression of SATB1 protein became hardly detectable by immunoblotting and its mRNA levels were substantially reduced by 70% and 90% in both *SATB1*-shRNA1 and *SATB1*-shRNA2 expressing cells, respectively (Fig. 2a). SATB1 expression remained unaltered in MDA-MB-231 cells expressing a control shRNA whose sequence did not match any known human gene.

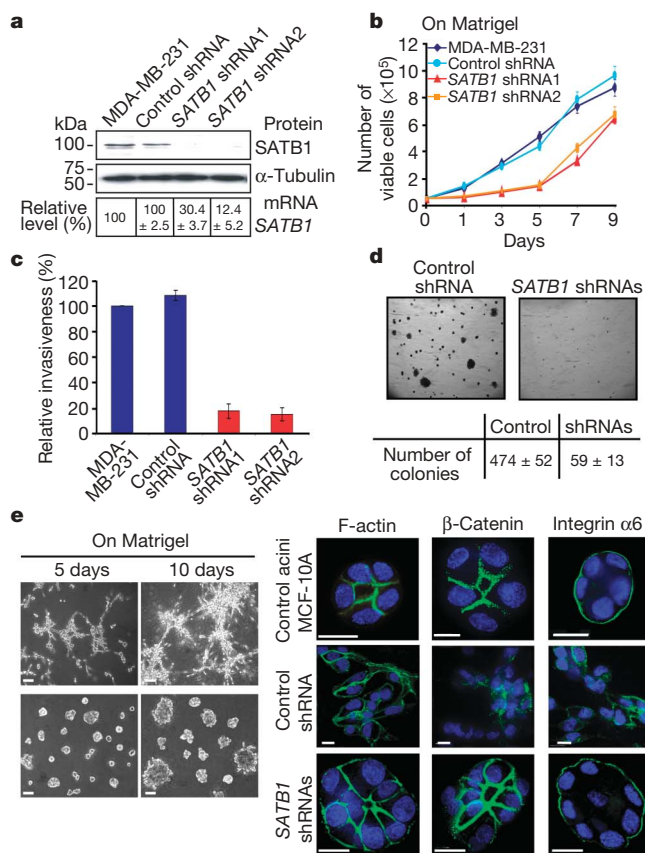
SATB1 knockdown decreased proliferation of *SATB1* shRNA1 and *SATB1* shRNA2 cells on Matrigel compared with the parental cell line and control shRNA cells (Fig. 2b). Furthermore, the invasive capacity *in vitro* of *SATB1* shRNA cells was reduced by 80–85% (Fig. 2c). Consistently, depletion of SATB1 prevented colony formation of these cell lines in soft agar, indicating that anchorage-dependent growth was restored (Fig. 2d).

Tissue organization and polarity are typically disrupted in mammary epithelial tumours *in vivo*. Therefore, we examined cell morphology of SATB1-depleted MDA-MB-231 cells. We observed major differences in cell morphology between control shRNA and *SATB1* shRNA1 cells grown on Matrigel. Where control cells exhibited a

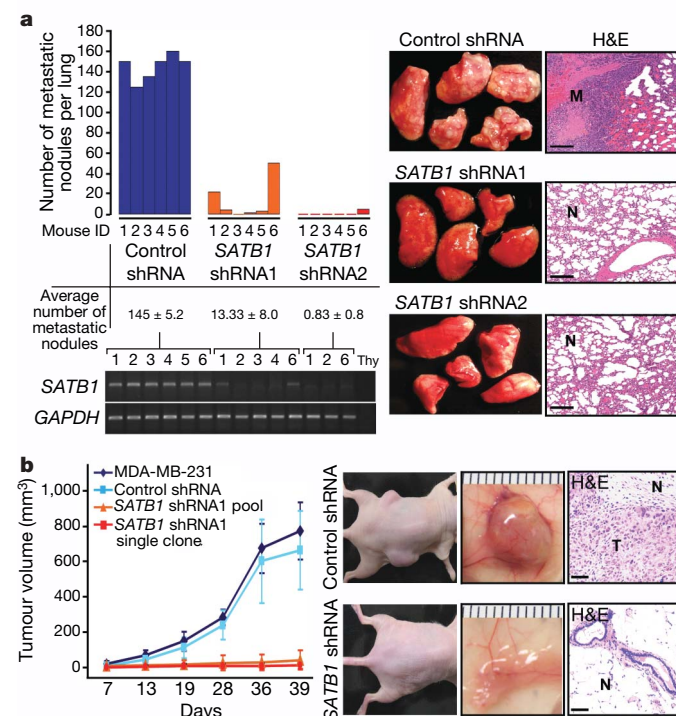
spindle-like fibroblastic morphology, *SATB1* shRNA cells had a cobble-stone-like morphology (Fig. 2e, left panel). We also observed similar cell morphology and reduction in invasive activity when SATB1 was depleted from the highly invasive BT549 breast tumour cell line (Supplementary Fig. 2a–d). When non-transformed mammary epithelial cells (MCF-10A cells) were cultured on Matrigel and analysed using markers of acinar formation—such as filamentous actin (F-actin),  $\beta$ -catenin, the basal extracellular matrix (ECM) receptor, and integrin  $\alpha 6$ —glandular-like structures (acini) formed, with a hollow lumen surrounded by polarized epithelial cells<sup>19–21</sup> (Fig. 2e, right panel). Control shRNA cells, however, formed disorganized structures lacking basal polarity (Fig. 2e, right panel). In contrast, *SATB1* shRNA1 cells cultured on Matrigel displayed normal acinar structures, showing uniform and polarized nuclei, cortically organized F-actin, basally distributed integrin  $\alpha 6$ , and  $\beta$ -catenin that localized to the lateral cell–cell junctions, as exhibited by MCF-10A cells. Similar results were confirmed with *SATB1* shRNA2 cells. Therefore, SATB1 knockdown in MDA-MB-231 breast cancer cells restores polarized cellular structures found in normal mammary epithelial cells.

### SATB1 depletion reverses cancer metastasis

We next evaluated the *in vivo* effects of SATB1 depletion on metastasis. We used an assay called experimental metastasis, in which we injected *SATB1* shRNA1, *SATB1* shRNA2, or control shRNA cells ( $1 \times 10^6$  cells) into the lateral tail vein of 6-week-old athymic mice



**Figure 2 | SATB1 depletion restores cell polarity and reduces aggressive phenotypes of MDA-MB-231 cells *in vitro*.** **a**, Reduced SATB1 expression, determined by immunoblot and quantitative RT–PCR analyses, in *SATB1* shRNA (1 and 2) cells, compared with controls (parental cell line MDA-MB-231 and control shRNA cells;  $\alpha$ -tubulin levels as loading control). **b**, Reduced proliferation of *SATB1* shRNA (1 and 2) cells grown on Matrigel compared with controls. **c**, Chemoinvasion assay of *SATB1* shRNA (1 and 2) cells, compared with controls (error bars indicate s.e.m.,  $n = 3$  experiments). **d**, Representative photographs of soft agar colony formation 20 days after culture of control shRNA and *SATB1* shRNA (1 and 2) cells, with mean colony counts from three dishes. **e**, Left panel: Phase contrast micrographs of control shRNA (top row) or *SATB1* shRNA1 cells (bottom row) cultured on Matrigel. Scale bar, 40  $\mu$ m. Right panel: morphologies of *SATB1* shRNA1 cells (acinar structure) and controls grown on Matrigel and stained for F-actin,  $\beta$ -catenin or integrin  $\alpha 6$  (green) and DAPI (blue). MCF-10A cells show the typical acinar structure. Scale bars, 15  $\mu$ m. SATB1 nuclear distribution in MDA-MB-231 cells is shown in Supplementary Fig. 2e.



**Figure 3 | SATB1 is necessary for lung colonization and tumour growth.** **a**, Number of metastases in lungs of mice ( $n = 6$  per group) 9 weeks after tail-vein injection of control shRNA, *SATB1* shRNA1 or *SATB1* shRNA2 cells (left), with mean nodule per lung values shown below. Expression levels of human *SATB1* in lungs was analysed by RT–PCR, with human *GAPDH* as a loading control and mouse thymocytes (Thy) as a negative control. Representative lungs and haematoxylin and eosin (H&E) staining of metastatic tumour (M) and normal (N) lung tissues are shown (right). **b**, Reduced tumour volumes in fat pads of nude mice injected with *SATB1* shRNA1 cells (single clones or pools), compared to controls (parental cell line MDA-MB-231 or control shRNA cells). Each data point is the mean value ( $\pm$  s.e.m.) of five–six primary tumours. Photographs of representative mice and tumours are shown, along with haematoxylin and eosin staining of tumour (T) and normal (N) breast tissue. Scale bar, 50  $\mu$ m.

and evaluated their survival in circulation and extravasation to and growth in the lung.

After 9 weeks, injected control tumour cells formed 125 to 160 metastatic nodules per lung in all 6 mice analysed, detected under a dissection microscope (Fig. 3a). In contrast, mice injected with *SATB1* shRNA1 cells formed 0 to 50 nodules per lung and mice injected with *SATB1* shRNA2 cells formed 0 to 5 nodules ( $n = 6$  mice each). Histological analyses confirmed that the number of micro-metastatic lesions was markedly reduced in the lungs of mice injected with *SATB1* shRNA1 or *SATB1* shRNA2 cells (Fig. 3a). The presence of human cancer cells in each lung was verified by polymerase chain reaction with reverse transcription (RT-PCR). Our data indicate that *SATB1* is necessary for the aggressive, highly metastatic phenotype of MDA-MB-231 cells.

We next tested whether *SATB1* depletion from MDA-MB-231 cells also inhibits tumour growth. We injected control shRNA and *SATB1* shRNA1 cells into the fourth mammary fat pads of athymic nude mice and monitored tumour growth. In contrast to both parental

MDA-MB-231 and control shRNA cells, which formed large tumours within 39 days (6 out of 6 mice), all six mice injected with an *SATB1* shRNA1 clone or a pool of *SATB1* shRNA1 cells resulted in either no tumours or greatly reduced tumour growth, respectively (Fig. 3b). These results indicate that *SATB1* expression in MDA-MB-231 cells is necessary for the tumour growth of these cells in mammary fat pads of mice.

### Promotion of growth and metastasis by *SATB1*

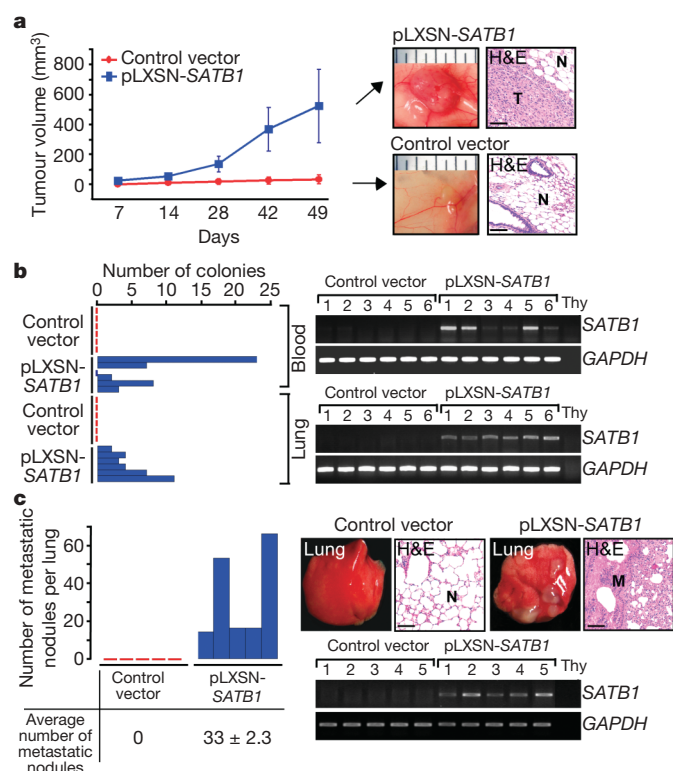
We examined whether ectopic expression of *SATB1* is sufficient to induce invasive activity in non-metastatic cancer cells. Control SKBR3 cells (a non-metastatic cell line, transfected with control vector) injected into the mammary glands of mice did not form tumours in mice after 7 weeks. In contrast, all six mice similarly injected with SKBR3 cells ectopically expressing *SATB1* (pLXSN-*SATB1*) grew large, undifferentiated, highly vascularized tumours (Fig. 4a). To examine intravasation, cells isolated from blood and lung tissue from mice injected with SKBR3 cells (control or pLXSN-*SATB1*) were cultured for 4 weeks in the presence of G418, to select for transfected cells. In 5 out of the 6 mice injected with pLXSN-*SATB1* cells, 2 to 23 colonies formed from each blood sample, and in all these mice 2–11 colonies formed from each lung sample. No colonies grew from samples from mice injected with control cells (Fig. 4b). These data show that expression of *SATB1* is sufficient to induce SKBR3 cells to form large tumours in mammary fat pads, to acquire the ability to invade blood vessels, and to survive in the circulation.

By 7 weeks after injection, we did not observe macroscopically visible metastases in the lungs of mice injected with *SATB1*-expressing cells in the mammary fat pads; longer monitoring times would be needed to observe such secondary tumours. However, we had to kill mice bearing large tumours after 7 weeks. Therefore, we intravenously injected SKBR3 cells (control vector or pLXSN-*SATB1*) into mice and found that at 10 weeks after injection, *SATB1*-expressing SKBR3 cells, but not control SKBR3 cells, formed many large metastatic nodules, indicating extravasation and tumour growth in lung (Fig. 4c).

We then analysed the invasive activity of breast cancer cells with differing levels of *SATB1* expression using Hs578T cells, a breast cancer cell line that endogenously expresses lower levels of *SATB1* than MDA-MB-231 cells. Compared to control Hs578T cells, Hs578T cells that overexpress *SATB1* (pLXSN-*SATB1* transfection; HS25) showed increased invasive activity *in vitro*, promoted tumour growth in mammary fat pads and displayed experimental metastasis to lung *in vivo* (Supplementary Fig. 3). These results indicate that high-level expression of *SATB1* is necessary and sufficient to promote metastasis of Hs578T cells to lung.

### Reprogramming of global expression profile

We performed gene expression profiling on MDA-MB-231 cells expressing either control shRNA or *SATB1* shRNA1 grown in culture. Unsupervised clustering analysis of 2,678 genes from two different microarray platforms (Affymetrix and Codelink) identified two groups of genes (tree 1 and tree 2) that significantly changed expression levels (by >1.5-fold) after *SATB1* depletion under both plastic and Matrigel culture conditions. Tree 1 included 409 down-regulated genes, and tree 2 contained 456 up-regulated genes on *SATB1* depletion (Fig. 5a). Functional profiling of these genes revealed that the greatest proportion of the genes was associated with cell adhesion, followed by phosphatidylinositol signalling and cell cycle regulation. Individual *SATB1*-dependent genes and subgroups of different molecular pathways are shown in Supplementary Tables 3 and 4 and Supplementary Fig. 4. Among the 231 Rosetta poor-prognosis-associated genes<sup>6</sup>, 174 were compared by our microarray to the *SATB1*-dependent gene set (shown in trees 1 and 2). Sixty-three of these genes (36%) whose expression was up- or down-regulated in breast tumours with poor prognoses were correspondingly altered by *SATB1* expression ( $P = 0.02$ ) (Supplementary Fig. 5a).



**Figure 4 | Ectopic expression of *SATB1* in SKBR3 cells induced tumour growth, intravasation and lung colonization.** **a**, Left panel: mean volumes ( $n = 6$  per group) of tumours formed in fat pads of mice injected with *SATB1*-overexpressing SKBR3 cells (pLXSN-*SATB1*) or controls (error bar indicates s.e.m.). Right panels: gross anatomy and haematoxylin-and-eosin-stained sections of tumours (pLXSN-*SATB1* cells) and normal breast tissues (control cells). T, tumour; N, normal breast tissue. Scale bar, 80 μm. **b**, Left panel: intravasation of tumour cells was determined from the numbers of colonies formed by pLXSN-*SATB1* or control cells grown in G418-containing media after isolation from blood or lung samples of tumour-bearing mice from experiments described in **a**. Right panel: RT-PCR analysis was used to detect human *SATB1* expression in cell colonies grown from blood and lung samples. *GAPDH* was the loading control and mouse thymocytes were the negative control (Thy). **c**, Left panel: the number of metastases that formed in the lungs of each nude mouse ( $n = 5$ ) 10 weeks after the injection of pLXSN-*SATB1* or control cells, and mean values for each group. Right panel: representative photos of the lungs and haematoxylin and eosin staining sections show normal lung from mice injected with control vector cells and metastatic nodules from pLXSN-*SATB1* cells. N, normal lung tissue; M, metastatic nodule. Lung nodules were shown to express human *SATB1* by RT-PCR. Scale bar, 80 μm.

Genes known to promote either bone<sup>22</sup> or lung metastasis<sup>23</sup> were also enriched among the SATB1-dependent genes in MDA-MB-231 cells ( $P = 0.0002$  and  $P = 0.021$ , respectively; Supplementary Fig. 5b, c).

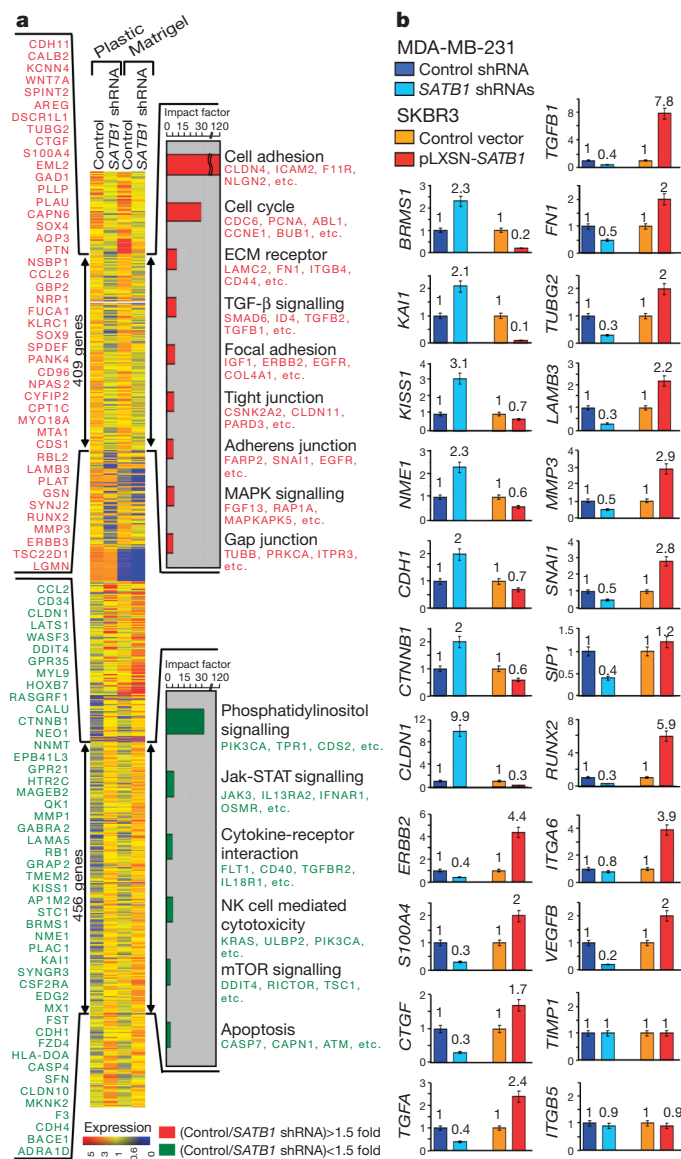
We used quantitative RT-PCR (qRT-PCR) and semi-qRT-PCR to confirm SATB1-dependent expression of over 40 genes identified in the microarray analysis (Fig. 5b and Supplementary Fig. 6). The expression of many genes known to have important functions in promoting metastasis was found to be upregulated by SATB1, including metastasin (*S100A4*)<sup>24</sup> and *VEGFB*<sup>25</sup>, which have roles in

metastasis and angiogenesis; matrix metalloproteases (MMPs) 2, 3 and 9, which degrade ECM and promote tumour invasion<sup>26,27</sup>; transforming growth factor- $\beta$ 1 (*TGFB1*), which stimulates invasion<sup>28</sup>; and connective tissue growth factor (*CTGF*), which mediates angiogenesis<sup>29</sup> and bone metastasis<sup>22</sup>. Many tumour cells exhibit increased invasiveness in response to TGF- $\beta$ 1 and increased levels of TGF- $\beta$ 1 have been reported in most tumour types<sup>28</sup>. Notably, SATB1 upregulates genes involved in epidermal growth factor (EGF) signalling<sup>30</sup>, such as the EGF receptor subfamily members *ERBB1*, *ERBB2* (*HER-2* or *NEU*), *ERBB3*, *ERBB4* and the ligands *NRG* and *AREG*. *ERBB2*, the most oncogenic member of the ERBB family, is an important regulator of breast cancer progression<sup>30</sup>, and drugs that intercept *ERBB2* signalling are in routine clinical application<sup>31</sup>. SATB1 also represses the metastasis suppressor genes *BRMS1*, *KAI1* (also called CD82) *KISS1* and *NME1* (*NM23*)<sup>32</sup>. SATB1 expression induces a marked change in the gene expression pattern in cancer cells and promotes their acquisition of aggressive phenotypes.

SATB1 depletion blocks the upregulation of cell-structure genes typical in invasive breast cancers (Fig. 5b and Supplementary Fig. 6), consistent with our observations that SATB1 depletion from MDA-MB-231 cells restores normal cell morphology. Such cell structure genes include the ECM protein fibronectin (*FN*), the intermediate filament protein vimentin (*VIM*)<sup>33</sup> and the cell-ECM interacting protein  $\beta$ 4 integrin (*ITGB4*)<sup>34</sup>. Dysregulated expression of cadherin and catenins, which mediate cell-to-cell adhesion, has been associated with breast cancer<sup>35,36</sup>. OB-cadherin (*CDH11*), VE-cadherin (*CDH5*) and N-cadherin (*CDH2*), often upregulated in invasive breast cancer, were all upregulated by SATB1. In contrast, genes downregulated by SATB1 included claudin 1 (*CLDN1*), a tight junction protein that is commonly lost or mislocalized in invasive tumours<sup>37</sup>;  $\beta$ -catenin (*CTNNB1*), a critical member of the canonical Wnt pathway<sup>38</sup>; and E-cadherin (*CDH1*), an adherens junction protein and tumour suppressor<sup>35,39</sup>. Loss of E-cadherin is a hallmark for epithelial to mesenchymal transition, a process whereby epithelial cells lose polarity, cell-to-cell contacts, and cytoskeletal integrity contributing to the dissemination of carcinoma cells from epithelial tumours<sup>40</sup>. On SATB1 depletion, the observed upregulation of E-cadherin and downregulation of fibronectin and repressors of E-cadherin such as *SNAI1* and *SIP1* indicate that the epithelial to mesenchymal transition process is reversed, resulting in the restoration of acinar-like morphology.

### SATB1 regulates gene activity and epigenetics

To identify genes directly regulated by SATB1, we determined the *in vivo* binding status of SATB1 within genomic loci of nine SATB1-dependent genes: *ERBB2*, *S100A4*, *ABL1*, *TGFB1*, *MMP3* and *LMNA* (laminA/C), which are all upregulated by SATB1, and *BRMS1*, *CLDN1* and *CTNNB1*, which are all downregulated by SATB1. *GAPDH*, *ITGB5* and *TIMP1* were selected as non-SATB1-dependent controls (Fig. 5b and Supplementary Fig. 6). For each of these 12 genes, we analysed a ~20-kb region upstream and downstream of the gene's first exon for potential SATB1 binding *in vivo*, looking for all potential SATB1 target sequences (BURs) (indicated by red number), promoter sequences (if known: blue box), regions containing CpG islands (green box), and other regulatory sequences (light blue box) (Fig. 6a–c and Supplementary Fig. 7a, b). Potential BURs were identified by the genomic sequences characterized by the ATC sequence context and confirmed for *in vitro* SATB1 binding by electrophoresis-mobility shift assay (EMSA, data not shown). To assess SATB1 binding to these loci *in vivo*, we used the urea-chromatin immunoprecipitation (urea-ChIP) method<sup>16</sup>. All genes analysed contained several BURs. Most BURs of SATB1-dependent genes were bound to SATB1, whereas none or only a few BURs in SATB1-independent genes were bound (Supplementary Fig. 7a and data not shown). The SATB1-binding status of SATB1-dependent or -independent genes indicates that SATB1 directly regulates expression of *ERBB2*, *S100A4*, *BRMS1*, *CLDN1*, *ABL1*, *TGFB1*, *LMNA*, *MMP3* and *CTNNB1*.



**Figure 5 | Global changes in expression profiles on SATB1 expression.** **a**, Unsupervised clustering (GeneSpring software) of genes differentially expressed between control shRNA and SATB1 shRNA1 cells from both plastic and Matrigel culture conditions; 409 SATB1-activated genes and 456 SATB1-repressed genes are marked by double-headed arrows. Representative SATB1-activated (red) and SATB1-repressed genes (green) are either listed vertically or under each molecular pathway. Impact factor strength of SATB1-activated (red bars) and -repressed (green bars) genes is shown. **b**, Expression levels of multiple genes in the microarray, including cancer-related genes, confirmed with quantitative RT-PCR: control shRNA (dark blue) versus a mixture of SATB1 shRNA1- and shRNA2-expressing clones grown on Matrigel (light blue); pooled SKBR3 control cells (orange) versus primary tumours from nude mice injected with SATB1-expressing cells (pLXSN-SATB1) (red). Genes analysed are shown on the y axis; numbers represent fold differences relative to *GAPDH*. Error bars indicate s.e.m. ( $n = 3$  experiments).

We addressed the mechanism by which SATB1 regulates the expression of its target genes. Using the urea-ChIP assay followed by qPCR (urea-ChIP-qPCR), we determined whether the histone acetylation status and the *in vivo* binding status of histone acetyltransferase (p300) and histone deacetylase 1 (HDAC1) were dependent on SATB1. We focused our analyses on *ERBB2*, *S100A4*, *BRMS1*, *CLDN1*, *ITGB5* and *TIMP1* (Fig. 6). In control cells (blue bars), the *in vivo* SATB1-binding sites in *ERBB2* and *S100A4* (*ERBB2*, positions 10, 12 and 13; *S100A4*, positions 2, 6, 7, 10, 12, 13 and 14) corresponded to the main acetylation peaks for histone H3 at lysines 9/14 (Fig. 6a). These positions coincide with p300 binding peaks *in vivo*. On SATB1 depletion (red bars), histone acetylation levels and p300 binding were markedly reduced. Instead, HDAC1 binding at these specific positions was increased. These epigenetic changes are

consistent with SATB1 promoting their gene expression. Opposite effects were observed for *BRMS1* and *CLDN1*, two genes repressed by SATB1 (Fig. 6b). On SATB1 depletion, histone H3 became more acetylated at SATB1-bound positions (*BRMS1*, positions 10 and 14; *CLDN1*, positions 9, 10 and 12), accompanied by increased p300 binding and reduced recruitment of HDAC1, leading to upregulation of these genes. Histone modifications and p300 binding status of non-target gene loci, transcriptionally active *ITGB5* and *TIMP1*, did not differ when SATB1 knockdown cells were compared with control MDA-MB-231 cells (Fig. 6c). Thus, SATB1 binds to its target gene loci at multiple positions and assembles the loci with histone-modifying factors to establish the epigenetic status.

## Discussion

We present a new model of gene regulation during tumour progression, in which the genome organizer SATB1 becomes expressed during malignancy, markedly altering the gene expression profile of breast cancer cells to induce an aggressive phenotype that promotes both tumour growth and metastasis.

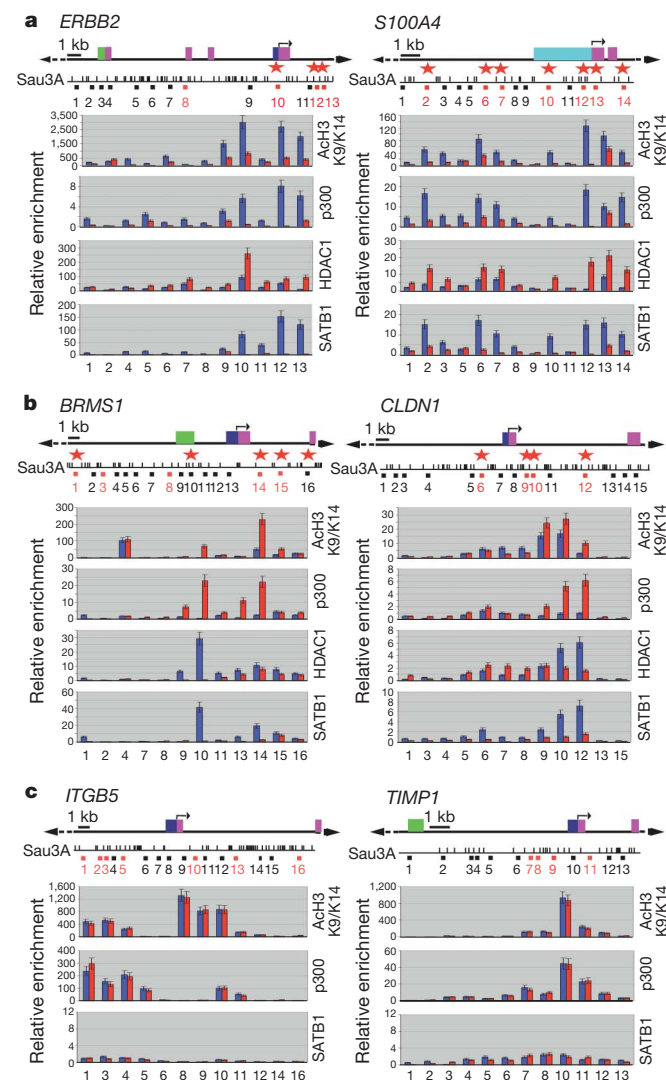
The expression of over one-thousand genes is altered on SATB1 expression in breast cancer cells. We found that SATB1 directly bound to and established the epigenetic status of all SATB1-dependent genes randomly chosen for analysis. There are likely to be a large number of additional genes regulated in a similar manner. An attractive mechanism by which SATB1 globally reprogrammes gene expression during metastasis is by tethering hundreds of gene loci onto its regulatory network, assembling them with chromatin modifying and transcription factors. Our findings support the emerging links between chromatin remodelling enzymes, epigenetics and cancer<sup>41–43</sup>. We have shown that a variety of genes involved in many aspects of tumorigenesis are regulated by SATB1, indicating that a large group of SATB1-targeted genes collectively induce tumour growth and metastasis. Similar to the cytokine genes in activated T cells<sup>18</sup>, it is likely that SATB1-target genes in breast cancer are brought together from intra- or even inter-chromosomal loci to form chromatin loops and are co-regulated by SATB1. These findings raise important questions regarding how tissue specificity of genes regulated by SATB1 is achieved (activated T cells versus breast cancer) and the mechanisms by which specific genes are either upregulated or repressed. Our results indicate that the SATB1-dependent gene expression pattern does not overlap completely between cancer cells cultured on plastic or on Matrigel, indicating crosstalk between the cellular environment and SATB1 nuclear architecture. It is possible that specific SATB1-regulated genes vary based on host organs to which cells metastasize. Other proteins that bind BURs<sup>44</sup>, such as HMGI(Y), which also has an important role in breast cancer progression<sup>45</sup>, may interact with SATB1 to promote cell growth in different cellular environments.

SATB1 expression is not restricted to late clinical stages of disease, but is observed in a subset of primary breast tumours at early clinical stages before lymph node metastasis. The SATB1 level in the nuclei of cancer cells has high prognostic significance, independent of the lymph node status ( $P < 0.0001$ ), indicating its utility in predicting the likelihood of disease progression in patients with early-stage breast cancer.

Although it is still essential to confirm the fate of SATB1-expressing cells in human primary breast tumours, our findings suggest a new paradigm in tumour progression, in which SATB1 functions as a genome organizer during tumorigenesis to reprogramme expression and promote metastasis. Future studies will address how applicable this concept is to other tissues. SATB1 may be useful as a therapeutic target for metastatic breast disease.

## METHODS SUMMARY

SATB1 expression levels were quantified by qPCR and immunoblot analyses with human breast cancer and non-malignant cell lines and with human tissue specimens from primary tumours and adjacent tissues. Prognostic values of



**Figure 6 | SATB1 defines the epigenetic status of target genes.** **a–c**, Urea-ChIP-qPCR was performed for MDA-MB-231 control shRNA (blue bar) and SATB1 shRNA (red bar) cells using antibodies against acetylated histone H3 at K9/14, p300, HDAC1 and SATB1. SATB1-upregulated genes *ERBB2* and *S100A4* (metastasin) (**a**), SATB1-downregulated genes *BRMS1* and *CLDN1* (**b**) and SATB1-independent genes *ITGB5* and *TIMP1* (**c**) were selected for the analysis. Gene structures are based on the data from USCS (<http://genome.ucsc.edu/>). Numbers indicate positions of DNA fragments where primers were designed (black, non-SATB1-recognizing sequences; red, SATB1-binding sequences shown by EMSA *in vitro*). Red stars indicate *in vivo* SATB1-bound DNA fragments. Dark blue box, promoter regions; light blue box, transcription binding sites; green box, CpG island; arrows, transcriptional start sites; pink box, exon. Error bars indicate s.e.m. ( $n = 3$  experiments).

SATB1 were evaluated by anti-SATB1 immunostaining of breast tissue microarrays containing >2,000 cases with clinical follow-up records. SATB1 expression in breast cancer cells was either knocked down by stable expression of shRNA against *SATB1* or induced by SATB1 expression from a retroviral expression construct, pLXSN-*SATB1*. Tumour growth and intravasation were studied by injecting *neo*-expressing human cancer cells proximal to the mammary glands in mice. Tumour growth was monitored for 6–7 weeks and the presence of human cells in blood and lung were determined by G418 selection. Tumour cells were intravenously injected in to mice and metastasis to lung was determined by quantifying lung colonization after 9–10 weeks. Colonized lung and mammary tumours were confirmed by pathological analyses. All animal work was done following Institutional Animal Care and Use Committee guidelines. Urea-ChIP-qPCR was performed to determine histone modification and binding status of modifying factors and SATB1 at selected gene loci. The expression and functional profiles of genes were compared between MDA-MB-231 cells expressing control shRNA and *SATB1* shRNA using Codelink Uniset Human 20K and Affymetrix HT-HG-U133A microarray chips (dataset on the GEO website; accession number GSE5417). Expression of ~40 genes was verified by qRT-PCR and semi-qRT-PCR.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 11 September 2007; accepted 22 January 2008.**

- Parker, B. & Sukumar, S. Distant metastasis in breast cancer: molecular mechanisms and therapeutic targets. *Cancer Biol. Ther.* **2**, 14–21 (2003).
- Chambers, A. F., Groom, A. C. & MacDonald, I. C. Dissemination and growth of cancer cells in metastatic sites. *Nature Rev. Cancer* **2**, 563–572 (2002).
- Fidler, I. J. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nature Rev. Cancer* **3**, 453–458 (2003).
- Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
- van de Vijver, M. J. *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**, 1999–2009 (2002).
- van't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536 (2002).
- Ince, T. A. & Weinberg, R. A. Functional genomics and the breast cancer problem. *Cancer Cell* **1**, 15–17 (2002).
- Sorlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl Acad. Sci. USA* **98**, 10869–10874 (2001).
- Ramaswamy, S., Ross, K. N., Lander, E. S. & Golub, T. R. A molecular signature of metastasis in primary solid tumors. *Nature Genet.* **33**, 49–54 (2003).
- Nguyen, D. X. & Massague, J. Genetic determinants of cancer metastasis. *Nature Rev. Genet.* **8**, 341–352 (2007).
- Alvarez, J. D. *et al.* The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. *Genes Dev.* **14**, 521–535 (2000).
- Dickinson, L. A., Joh, T., Kohwi, Y. & Kohwi-Shigematsu, T. A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell* **70**, 631–645 (1992).
- Yasui, D., Miyano, M., Cai, S., Varga-Weisz, P. & Kohwi-Shigematsu, T. SATB1 targets chromatin remodelling to regulate genes over long distances. *Nature* **419**, 641–645 (2002).
- Cai, S., Han, H. J. & Kohwi-Shigematsu, T. Tissue-specific nuclear architecture and gene expression regulated by SATB1. *Nature Genet.* **34**, 42–51 (2003).
- Kohwi-Shigematsu, T. & Kohwi, Y. Torsional stress stabilizes extended base unpairing in suppressor sites flanking immunoglobulin heavy chain enhancer. *Biochemistry* **29**, 9551–9560 (1990).
- Kohwi-Shigematsu, T., deBelle, I., Dickinson, L. A., Galande, S. & Kohwi, Y. Identification of base-unpairing region-binding proteins and characterization of their *in vivo* binding sequences. *Methods Cell Biol.* **53**, 323–354 (1998).
- Bode, J. *et al.* Biological significance of unwinding capability of nuclear matrix-associating DNAs. *Science* **255**, 195–197 (1992).
- Cai, S., Lee, C. C. & Kohwi-Shigematsu, T. SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nature Genet.* **38**, 1278–1288 (2006).
- Debnath, J. *et al.* The role of apoptosis in creating and maintaining luminal space within normal and oncogene-expressing mammary acini. *Cell* **111**, 29–40 (2002).
- Weaver, V. M. *et al.*  $\beta 4$  integrin-dependent formation of polarized three-dimensional architecture confers resistance to apoptosis in normal and malignant mammary epithelium. *Cancer Cell* **2**, 205–216 (2002).
- Underwood, J. M. *et al.* The ultrastructure of MCF-10A acini. *J. Cell. Physiol.* **208**, 141–148 (2006).
- Kang, Y. *et al.* A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* **3**, 537–549 (2003).
- Minn, A. J. *et al.* Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518–524 (2005).
- Helfman, D. M., Kim, E. J., Lukanidin, E. & Grigorian, M. The metastasis associated protein S100A4: role in tumour progression and metastasis. *Br. J. Cancer* **92**, 1955–1958 (2005).
- Salven, P. *et al.* Vascular endothelial growth factors VEGF-B and VEGF-C are expressed in human tumors. *Am. J. Pathol.* **153**, 103–108 (1998).
- Chang, C. & Werb, Z. The many faces of metalloproteinases: cell growth, invasion, angiogenesis and metastasis. *Trends Cell Biol.* **11**, S37–S43 (2001).
- Duffy, M. J., Maguire, T. M., Hill, A., McDermott, E. & O'Higgins, N. Metalloproteinases: role in breast carcinogenesis, invasion and metastasis. *Breast Cancer Res.* **2**, 252–257 (2000).
- Dumont, N. & Arteaga, C. L. Targeting the TGF $\beta$  signaling network in human neoplasia. *Cancer Cell* **3**, 531–536 (2003).
- Moussad, E. E. & Briggstock, D. R. Connective tissue growth factor: what's in a name? *Mol. Genet. Metab.* **71**, 276–292 (2000).
- Mosesson, Y. & Yarden, Y. Oncogenic growth factor receptors: implications for signal transduction therapy. *Semin. Cancer Biol.* **14**, 262–270 (2004).
- de Bono, J. S. & Rowinsky, E. K. The ErbB receptor family: a therapeutic target for cancer. *Trends Mol. Med.* **8**, S19–S26 (2002).
- Steege, P. S. Metastasis suppressors alter the signal transduction of cancer cells. *Nature Rev. Cancer* **3**, 55–63 (2003).
- Kang, Y. & Massague, J. Epithelial-mesenchymal transitions: twist in development and metastasis. *Cell* **118**, 277–279 (2004).
- Vidal, F. *et al.* Integrin  $\beta 4$  mutations associated with junctional epidermolysis bullosa with pyloric atresia. *Nature Genet.* **10**, 229–234 (1995).
- Takeichi, M. Cadherins in cancer: implications for invasion and metastasis. *Curr. Opin. Cell Biol.* **5**, 806–811 (1993).
- Cowin, P., Rowlands, T. M. & Hatsell, S. J. Cadherins and catenins in breast cancer. *Curr. Opin. Cell Biol.* **17**, 499–508 (2005).
- Tokes, A. M. *et al.* Claudin-1, -3 and -4 proteins and mRNA expression in benign and malignant breast lesions: a research study. *Breast Cancer Res.* **7**, R296–R305 (2005).
- Brembeck, F. H., Rosario, M. & Birchmeier, W. Balancing cell adhesion and Wnt signaling, the key role of  $\beta$ -catenin. *Curr. Opin. Genet. Dev.* **16**, 51–59 (2006).
- Berx, G. & Van Roy, F. The E-cadherin/catenin complex: an important gatekeeper in breast cancer tumorigenesis and malignant progression. *Breast Cancer Res.* **3**, 289–293 (2001).
- Thiery, J. P. Epithelial-mesenchymal transitions in tumour progression. *Nature Rev. Cancer* **2**, 442–454 (2002).
- Roberts, C. W. & Orkin, S. H. The SWI/SNF complex—chromatin and cancer. *Nature Rev. Cancer* **4**, 133–142 (2004).
- Baylin, S. B. & Ohm, J. E. Epigenetic gene silencing in cancer—a mechanism for early oncogenic pathway addiction? *Nature Rev. Cancer* **6**, 107–116 (2006).
- Drobic, B., Dunn, K. L., Espino, P. S. & Davie, J. R. Abnormalities of chromatin in tumor cells. *EXS* **96**, 25–47 (2006).
- Galante, S. & Kohwi-Shigematsu, T. Linking chromatin architecture to cellular phenotype: BUR-binding proteins in cancer. *J. Cell. Biochem., Suppl.* **35**, 36–45 (2000).
- Reeves, R., Edberg, D. D. & Li, Y. Architectural transcription factor HMGI(Y) promotes tumor progression and mesenchymal transition of human epithelial cells. *Mol. Cell. Biol.* **21**, 575–594 (2001).
- Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515–527 (2006).
- Holst, F. *et al.* Estrogen receptor  $\alpha$  (ESR1) gene amplification is frequent in breast cancer. *Nature Genet.* **39**, 655–660 (2007).
- Severgnini, M. *et al.* Strategies for comparing gene expression profiles from different microarray platforms: application to a case-control experiment. *Anal. Biochem.* **353**, 43–56 (2006).
- Draghici, S., Khatri, P., Martins, R. P., Ostermeier, G. C. & Krawetz, S. A. Global functional profiling of gene expression. *Genomics* **81**, 98–104 (2003).
- Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. F. & Fraser, P. Long-range chromatin regulatory interactions *in vivo*. *Nature Genet.* **32**, 623–626 (2002).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank J. W. Gray and M. Stamfers for providing some of the cell lines, M. J. Bissell, C. W. Roberts, J. A. Nickerson and S. A. Krauss for critical reading of the manuscript and useful suggestions, K. Novak and M. Kohwi for help in manuscript preparation, and R. Simon and M. Falduto for assisting expression microarray data analysis. This work was supported by a National Institute of Health grant to T.K.-S. and also by University of California Breast Cancer Research Program at its initial stage.

**Author Information** The expression data set is on the GEO website under accession number GSE5417. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to T.K.-S. (Terumiks@lbl.gov) or Y.K. (YKohwi@lbl.gov).

## METHODS

**Cell culture.** Breast cancer cell lines were obtained from American Type Culture Collection (ATCC) and maintained as described<sup>46</sup>. Immortalized mammary epithelial cell lines 184A1, 184AA2 and 184V were obtained from M. Stammers. Normal human mammary epithelial cells (Cambrex) and the retroviral packaging cell line PT67 (Clontech) were maintained according to the manufacturer's instruction. To obtain primary tumour cell cultures, tumour samples from nude mice were subjected to mechanical and enzymatic dissociation. The resulting cancer cells were cultured in DMEM containing G418 or puromycin. Protein and RNA samples were extracted from subconfluent cells in the exponential phase of growth.

**Human breast cancer specimens.** Human primary breast carcinomas, benign breast lesions, and normal breast tissues surgically removed and snap frozen in liquid nitrogen were obtained from the Cooperative Human Tissue Network (CHTN) along with pathology summaries. Protein extracts were prepared from 10 normal tissues and 28 tumour tissues. All human tissue samples were obtained and handled in accordance with an approved Institutional Review Board application.

**Cell growth assay.** *In vitro* proliferation was measured by seeding approximately  $5 \times 10^4$  cells on Matrigel-coated 24-well plates as previously described<sup>20</sup>. At specific time points, cells were isolated by incubation with dispase (BD Biosciences) for 2 h at 37 °C and then with trypsin for 5 min before counting. Trypan blue exclusion analysis indicated that 99–100% of the cells were viable.

**Chemoinvasion assay.** Assays were performed in 24-well chemotaxis plates with an 8 µm polycarbonate filter membrane coated with growth-factor-reduced Matrigel diluted in the range of 10% to 25%. Breast cancer cells in serum-free medium ( $2.5 \times 10^4$  cells per well) were added to the upper chamber and conditioned media derived from NIH3T3 fibroblast cultures was placed in the lower chambers as a chemo-attractant. The chambers were incubated for 20 h at 37 °C with 5% CO<sub>2</sub>; experiments were performed in triplicate. Migrated cells on the undersides of filter membrane were then fixed in 10% (w/v) buffered formalin and stained with crystal violet. The migrated cells were counted using light microscopy and s.e.m. values were determined for each sample.

**Soft agar assay.** Cells ( $1 \times 10^4$ ) were resuspended in DMEM containing 5% FBS with 0.3% agarose and layered on top of 0.5% agarose in DMEM on 60-mm plates. Cultures were maintained for 20 days. Colonies that grew beyond 50 µm in diameter were scored as positive. Each experiment was done in triplicate.

**Analysis of mRNA and protein expression.** Total RNA was purified using TRI reagent (Sigma) and the RNeasy kit (Qiagen) and then 5 µg of each sample was reverse transcribed using the SuperScript II RNaseH first-strand synthesis system (Invitrogen). cDNAs were analysed, in triplicate, using an ABI 7500 Fast Real-Time PCR System (Applied Biosystem). Semi-qRT-PCR was performed as previously described<sup>14</sup>. Protein expression levels were assessed by immunoblot analysis with cell lysates (40–60 µg) in lysis buffer (20 mM HEPES (pH 7.9), 25% glycerol, 0.5 N NaCl, 1 mM EDTA, 1% NP-40, 0.5 mM dithiothreitol, 0.1% deoxycholate) containing the protease inhibitors (Roche) using anti-SATB1 (BD Bioscience) and anti- $\alpha$ -tubulin antibodies (Sigma). Primer sequences used in RT-PCR experiments are available in Supplementary Table 5.

**Tissue microarray slides and immunohistochemistry.** Immunohistochemistry was performed using a peroxidase detection system with human breast cancer tissue microarray slides (TriStar), as previously described<sup>47</sup>. Rabbit polyclonal SATB1 antibody (1583) was pre-absorbed against SKBR3 cell lysates fixed on activated PVDF membrane, applied (1:1,800) and the slides were incubated overnight at 4 °C. The slides were then counterstained with haematoxylin and mounted with permount (Fisher). To evaluate SATB1 levels, immunostained slides were scored with digital images obtained by the ScanScope XT system (Aperio). The signal was scored based on the intensity and percentage of cells with SATB1 nuclear staining on the following scale: score 0, negative nuclear staining for all tumour cells; score 1, weak nuclear staining representing all positive staining other than score 2; score 2, moderate nuclear staining >50% or strong nuclear staining in >5% of the tumour cells. Samples that could not be interpreted or were missing most of the tumour tissue were given a score of not applicable (N/A). Scoring of the tissue microarray was completed by three independent observers. Significance of correlation between SATB1 signal and histopathological factors was determined using Pearson's chi-squared ( $\chi^2$ ) test. Kaplan–Meier plots were used to estimate the prognostic relevance of SATB1 in univariate analysis using WinStat (Fitch Software). Multivariate analysis was performed applying COX proportional hazards test.

**Immunofluorescence analysis.** Cultured cells were fixed in 4% paraformaldehyde, permeabilized in 0.1% Triton-X100 and blocked in 5% BSA. Focal adhesion complexes were detected with anti- $\beta$  catenin (clone 14) and anti-integrin  $\alpha$ 6 (CD49f; BD Biosciences). F-actin was detected with fluorescent phalloidin (Invitrogen). To detect epithelial cells, tissue microarray sections were stained

with anti-CK8 or anti-CK14 (Lab Vision) overnight at 4 °C, followed by Alexa Fluor 488 and/or Alexa Fluor 594 secondary antibodies (Invitrogen). Images were collected using a Delta Vision microscope and processed with SoftWoRx software (Applied Precision).

**SATB1-knockdown cells.** Two shRNAs were designed, based on the SATB1 sequence (NM\_002971) identified with siRNA Target Finder (Ambion): shRNA<sub>2176</sub> 5'-GGATTGGAAGAGAGTGTC-3', or shRNA<sub>2566</sub> 5'-GTCCACC-TTGTCTCTCTC-3'. The oligoduplexes were cloned into pSUPER-puro (Oligoengine), and transfected into cells using Lipofectamine 2000 (Invitrogen). Twenty-four hours later, transfected cells were selected for 10 days with 2 µg ml<sup>-1</sup> puromycin. Cells stably expressing shRNA<sub>2176</sub> or shRNA<sub>2566</sub> were designated SATB1 shRNA1 or SATB1 shRNA2, respectively. Pooled populations of knockdown cells, obtained after 10 days of drug selection without subcloning, were injected into nude mice for *in vivo* experiments. Negative control cell lines were generated by infecting cells with a pSUPER-puro construct targeting EGFP cDNA (5'-GAAGCAGCAGCACTTCTC-3') which did not yield any appreciable knockdown of the protein product in immunoblot analysis.

**SATB1-overexpressing cells.** The human SATB1 cDNA was cloned into the pLXSN retroviral expression vector (Clontech) and then transfected into the PT67 packaging cell line using FuGene (Roche). Stable cell lines were selected by incubation with 1 mg ml<sup>-1</sup> G418. Virus-containing supernatants from PT67 cells were collected after 48–96 h incubation, concentrated, and the titre was checked. Viral media was added to 70%-confluent cells in the presence of 8 µg ml<sup>-1</sup> polybrene (Sigma) and incubated overnight. Forty-eight hours later, cells were incubated with 600 µg ml<sup>-1</sup> of G418.

**Expression microarray analysis.** Two microarray platforms were used: Codelink Human Uniset 20K (GE healthcare) and high throughput array (HTA) genechip HT-HG-U133A (Affymetrix). Samples were prepared according to the manufacturer's instructions and previous publications<sup>46</sup>. Initial analysis of microarray data was performed using CodeLink expression analysis (GE Healthcare) and GeneSpring software (Silicon Genetics). For cross-platform comparison, genes represented on both microarrays were identified by comparing annotated probe sets with Affymetrix's NetAffix Analysis Center and Codelink iCenter. Codelink systematic IDs were obtained for 21,080 of 22,215 (95%) Affymetrix GeneChip probe sets by a homology search using GeneSpring software. The strategies were applied to generate gene lists from different platforms as described previously<sup>48</sup>. For sample clustering, standard correlation was applied to measure the similarity of the expression pattern between different samples. The web-based software tools Onto-Express and Pathway-Express (<http://vortex.cs.wayne.edu/>) were used for functional profiling on the basis of gene ontology terms with obtained data<sup>49</sup>.

**Analysis of metastasis.** Six female NCR athymic mice were injected with MDA-MB-231 control cells (group 1) or SATB1 shRNA1 cells (group 2), five were injected with SKBR3 control (group 3) or SATB1-overexpressing SKBR3 cells (group 4), five were injected with Hs578T control cells (group 5), and six were injected with SATB1-overexpressing Hs578T cells (group 6). Cells ( $1 \times 10^6$  for groups 1–4;  $2 \times 10^6$  for groups 5 and 6) were injected intravenously, via the lateral tail vein, in 100 µl PBS. At 9–10 weeks after injection, when mice had not died but some appeared to be sick, all mice were killed and their lungs were removed and fixed in 10% formalin. The number of surface metastases per lung was determined under a dissecting microscope. Histopathological analyses were performed by the mutant mouse histopathology laboratory at University of California, Davis.

**Analysis of tumour growth and intravasation.** Six female NCR athymic mice were injected with: a pool of SATB1 shRNA1 cells (group 1); a single SATB1 shRNA1 clone with no detectable SATB1 (group 2); parental MD-MB-231 cells (group 3); control shRNA cells (group 4); SKBR3 control cells (group 5); SATB1-overexpressing SKBR3 cells (group 6); Hs578T control cells (group 7); or SATB1-overexpressing Hs578T cells (group 8). Cells were injected into the fourth mammary fat pad from flank ( $2 \times 10^5$  cells with Matrigel at 5 mg ml<sup>-1</sup> of PBS in a volume of 200 µl). Tumour growth was monitored externally using vernier calipers for 6–7 weeks. To examine the presence of human tumour cells, lung tissue (one half of lung) and 100 µl of blood were collected from the mice injected with SKBR3 cells (groups 5 and 6) immediately after they were killed; samples were incubated with 0.2% collagenase type-2 in DMEM for 2 h at 37 °C; and cells were dispersed, washed and plated in DMEM media plus G418 (600 µg ml<sup>-1</sup>). The tumour colonies were counted 4 weeks later.

**Urea-ChIP-qPCR assay.** Urea-ChIP experiments were carried out as previously described<sup>18</sup>. Briefly, formaldehyde-crosslinked chromatin from MDA-MB-231 control shRNA and SATB1 shRNA1 cells was purified by urea-gradient ultracentrifugation, digested with 60 U of Sau3A1, and resulting chromatin fragments were immunoprecipitated with anti-SATB1 (BD Bioscience), anti-acetylated H3-Lys9/Lys14, anti-p300 (Upstate), or anti-HDAC1 (Santa Cruz) antibodies. Purified mouse IgG1 (Sigma) was used as a control. Immunoprecipitated DNA

and input DNA were quantified using ND-1000 Spectrophotometer (NanoDrop Technologies). Quantitative ChIP-PCR was performed using platinum SYBR green qPCR supermix-UDG kit (Invitrogen) using an ABI 7500 Fast Real-Time PCR System (Applied Biosystem); all experiments were performed three times. The primer sequences were designed against the promoter regions of each gene, to cover ~20 kb, using Vector NTI software (Invitrogen); sequences are available on request. The relative enrichment of ChIP DNAs was determined by absolute quantification method and the formulas as previously described<sup>50</sup>. The gel mobility shift assay (EMSA) was performed to confirm the SATB1 binding sites *in vivo*. For semi-quantitative ChIP-PCR, 10 ng of reverse crosslinked chromatin was equally aliquoted into tubes, and performed PCR reaction using GeneAmp PCR system 9700 (Applied Biosystem) at 35–45 cycles. Primer sequences used in ChIP-qPCR and ChIP-semi-qPCR experiments are available in Supplementary Table 5.